REMARKS

Overview of amendments

The independent claims have been amended to correct an error in the original claim language, which originally recited a "depth image production section." The claims have now been amended to recite "depth information production section."

The independent claims are further amended to clarify the previous claim language. The previous claim language stated that images are converted to so that their pixel units represent an equal amount of an object. The claims are now amended to clarify that the images are converted so that the resolutions of each image coincide with each other. This language is used, for example, in the final paragraph on page 15 of the application.

The dependent claims are amended to use language that is consistent with the amended independent claims.

New claim 19 is a method claim that generally corresponds to the features recited in claim 2.

No new matter is added.

Response to rejection

The 15 July 2003 official action withdraws this application from appeal, removes the Subbarao reference from the obviousness rejections of all claims, and asserts the new rejection that all claims are obvious over the previously cited Auty reference in view of the admitted prior art. Auty is cited as teaching the features not taught by the admitted prior art, namely, the conversion of the pixel units of images used for distance measurement by triangulation. The claims now recite this feature in different language than was addressed by the official action, however the underlying concept is essentially the same, and so the present remarks will address the same portions of Auty that were cited as teaching these features.

Applicant sincerely respects the examiner's position that the features not taught by the admitted prior art are found in Auty. However, applicant has reviewed Auty many times and continues to be unable to understand why the examiner believes that Auty teaches these features. Applicant would like to resolve this conflict and come to an agreement as to what is taught by Auty. Accordingly, in this reply applicant will set out in detail his understanding of the portions of Auty that are asserted to teach this feature. The undersigned will then contact the examiner after filing this reply to schedule a telephone conference to discuss the points set out here and to see if an agreement can be reached.

Teaching of Auty

The official action asserts that Auty, at Figure 17 and col. 20, line 29 - col. 21, line 63, teaches conversion of the pixel units of images that are used for performing a distance measurement.

Applicant does not have the same understanding. Applicant believes that this portion of Auty teaches a process in which the position of a vehicle relative to the field of a main camera is determined from its location within the wider image field of a separate detection camera. Below applicant has reproduced the cited portions of Auty as well as additional portions of Auty to explain why applicant disagrees with the examiner's position and holds this contrary belief.

The purpose of Auty's system is to photograph vehicles on the highway for license plate identification. The photographs of the vehicles are taken by a main camera mounted in a fixed position over the highway. An issue faced by this system is determining when a vehicle is within the field of view of the main camera so that its photograph can be taken. To solve this problem, Auty uses a "detection" camera that is located at a fixed position near the main camera. The detection camera is a wide angle camera with a field of view that sees down the road beyond the field of the main camera. Auty's system tracks the movement of vehicles toward the field of the main camera by following their movement through the field of the detection camera. When it is determined

through this tracking that a vehicle is in range of the main camera, the main camera is triggered to take a photograph.

Figure 3 of the patent (below) illustrates how the cameras work together.

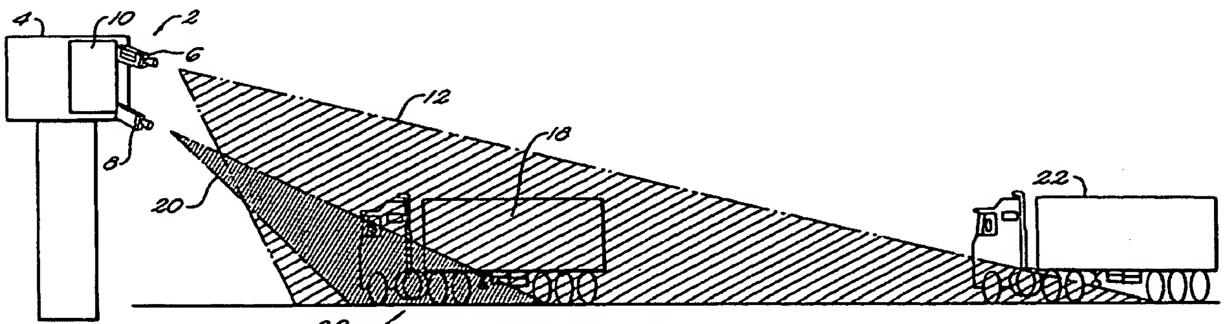

FIG. 3

Figure 3 shows the long field 12 of the detection camera 6 that is used to track vehicles, and the small field 20 of the main camera 8 that takes the actual vehicle photographs. Auty's system tracks the movement of a vehicle 22 using the detection camera 6 to determine when to take its photograph using the main camera 8.

Since there is only one detection camera for tracking the vehicle, its location on the road must be determined from its location within the image made by the single detection camera. Figure 9 shows an example of an image produced by the detection camera:

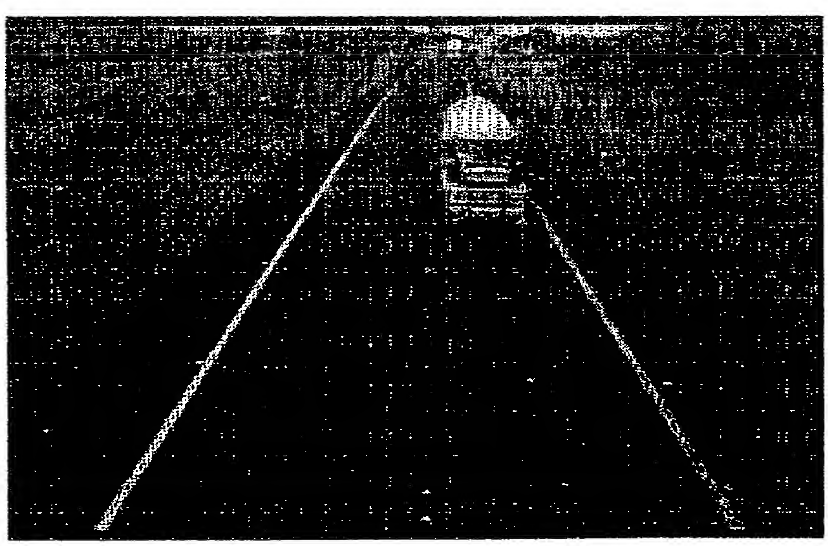

FIG. 9

This image does not directly tell the system where the vehicle is located. Therefore the system must compute that location by converting the two-dimensional coordinates of the vehicle within the detector camera image plane to three-dimensional real-world road coordinates. This conversion can be done directly because the detection camera is in a fixed position relative to the road, and so a vehicle at a given "height" in the detection camera image plane must be at a given location along the road. The following mark-up of Figure 9 shows the basic principle behind this conversion:

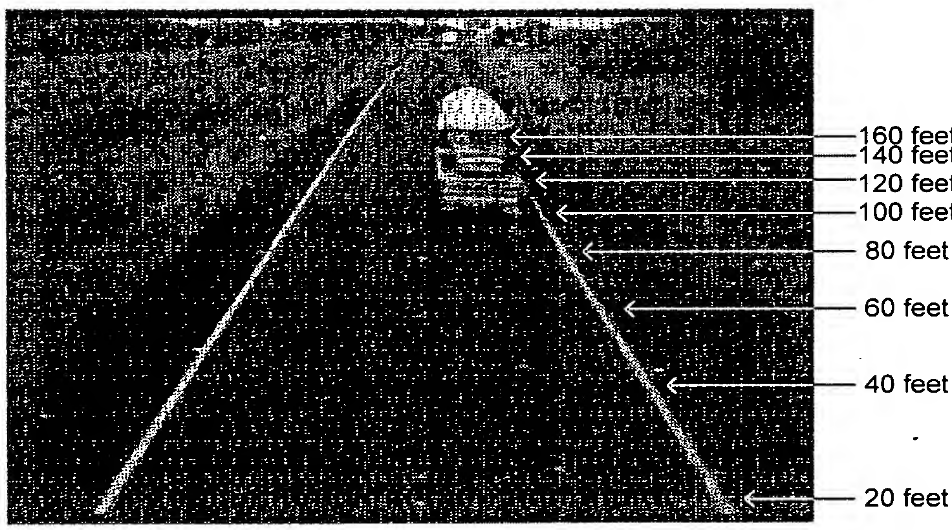

FIG. 9

Because the detection camera is in a fixed location relative to the road, an object that is at a given vertical position in the image will necessarily be at a given position along the road. Therefore the computer system performs a conversion that determines the position of the vehicle along the "y-axis" of the road based on its position in the detection camera image. The y-axis of the road is shown in Figure 25:

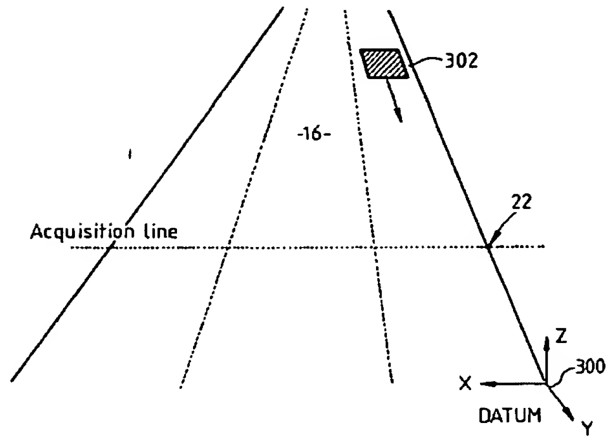

FIG. 25

By converting the location of the vehicle in the image plane to a location along the y-axis (the road), Auty determines where the vehicle is with respect to the "acquisition line," i.e. the line marking the location at which the main camera is triggered to photograph the vehicle. This is the conversion that is discussed in the disputed passage at col. 20, lines 29 to col. 21, line 63. That passage is reproduced below, along with applicant's notes indicating applicant's understanding of that passage:

The trajectory task 180 uses the received cluster data to track the clusters over successive video fields. The program file which controls the operation is traj.c. The coordinates used for tracking a cluster box are the coordinates of the centre of the base of the box, and the coordinate system for the roadway 16 which is adopted is illustrated in FIG. 25. The datum 300 of the roadway coordinate system is an arbitrary point on the roadway, which has been chosen as the centre of the left hand fog line underneath the edge of an overpass bridge holding the cameras 6 and 8. Vehicles 302 travel in positive Y axis direction on the roadway 16, starting at a negative value in the distance. The trajectory of a cluster box in image plane coordinates (x_i, y_i) as shown in the graph of FIG. 26 is not linear with time due to the effect of perspective. Therefore a camera transformation is applied so as to convert image plane coordinates to real world 3-D coordinates. In matrix form, the overall camera transformation is as follows:

This section explains that the location of a vehicle (cluster box) in the detector camera image is converted to a "real world 3-D" location along the road

$$\begin{bmatrix} x^i \\ y^i \\ z^i \end{bmatrix} = \begin{bmatrix} a_x & 0 & X_0 & 0 \\ 0 & a_y & Y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/f & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} {}^{0}T_{CAM}^{-1} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (13)$$

where

α_x	X-axis scaling factor in pixels/mm (intrinsic)
α_y	Y-axis scaling factor in pixels/mm (intrinsic)
X_0	image plane offset in pixels (intrinsic)
Y_0	image plane offset in pixels (intrinsic)
f	focal length (intrinsic)
${}^{0}T_{CAM}$	detection camera 6 position in world coordinates (extrinsic)

This section shows the matrix calculation that relates image plane coordinates (left) to real world coordinates (right-most). The terms in the center are fixed characteristics of the detection camera and its location.

The portion shown above gives the basic conversion equation. There are three important points to note about this conversion:

1) The conversion does not involve anything relating to the main camera. None of the terms of the equation are related to the main camera or its resolution. If the conversion was done with respect to a feature of the main camera then the conversion equation would include terms representing a relationship to those features of the main camera. The conversion equation does not include such terms.

2) The conversion does not involve converting the resolution of the detector image to match the resolution of another image. If the conversion was for the purpose of converting the resolution of the detector image, there would need to be terms in the conversion equation that represent a relationship between the resolution of the detector camera and the resolution of another camera. The conversion equation does not include such terms.

3) The reference to number of pixels is not used to convert detector camera images so that their pixels represent equal amounts of object in a main camera image. The X-axis and Y-axis scaling factors that relate a detector

image plane distance (a number of pixels) to a real-world distance in real-world units (mm) convert a location in the image to a location in the real world.

The specification goes on as follows:

The intrinsic parameters are innate characteristics of the camera and sensor, while the extrinsic parameters are characteristics only of the position and orientation of the camera. The principle point of the image plane is the intersection of the optical axis and that plane, at coordinates (X_0, Y_0). Equation 13 can be written as:

21

$$\begin{bmatrix} x^1 \\ y^1 \\ z^1 \end{bmatrix} = C \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{14}$$

where C is the camera calibration matrix, a 3×4 homogeneous transform which performs scaling, translation and perspective correction. The image plane coordinates are then expressed in terms of homogeneous coordinates as:

$$X^i = \frac{x^i}{z^i} \tag{15}$$

$$Y^i = \frac{y^i}{z^i} \tag{16}$$

This section reiterates that the conversion converts a real-world location to an image plane location.

The general perspective transform maps a ray in three dimensional space to a point on the image plane. For vehicle coordinates in the image plane as seen by the detection camera 6, a unique three dimensional location of the vehicle cannot be determined so the bottom of a cluster box received from the label task is considered to be on the roadway, i.e. z=0, and therefore the box can be tracked with reference to the roadway x and y coordinates. The equation 14, 15 and 16, given the image plane coordinates and z, can be solved simultaneously for the roadway coordinates x and y to specify the position of a vehicle. The equations have been solved using the computer algebra package MAPLE, and the solution, in C notation, is as follows:

```
den = (-Xi*C31*C22 + Xi*C32*C21 +
        (Yi*C21 - C21)*C12 + (-Yi*C32 + C22)*C11);
y = -(-Xi*C31*C24 + Xi*C34*C21 + (Yi*C21 - C21)*C14 +
      (Xi*C33*C21 - Xi*C31*C23)*z + (Yi*C31 - C21*z*C13 +
        (-yI*C34 + C24 + (-Yi*C33 + C23)*z)*C11)/den;
x = (-C24*Xi*C32 + C22*Xi*C34 + (Yi*C32 - C22)*C14 +
      (C22*Xi*C33 - C23*Xi*C32)*z + (Yi*C32 - C22)*z*C13 +
        (-Yi*C34 + C24 + (-Yi*C33 + C23)*z)*C12)/den;
```

The solution explicitly includes height above the roadway, z, which can be set at zero for daytime operation or some marginal distance above the roadway, whereas at night, the bottom of the cluster box generally corresponds to the height of the headlights above the road, and therefore z is set to a national headlight height. FIG. 27 illustrates a graph of the same vehicle trajectory as in FIG. 26, after the trajectory has been mapped to the roadway coordinates x and y. The trajectory illustrates the vehicle is moving at a constant speed, and in the left hand lane.

This section discusses the vehicle trajectory information that is generated by performing the conversion. See below.

The last paragraph above refers to the results shown in Figure 26 and Figure 27, which are reproduced below:

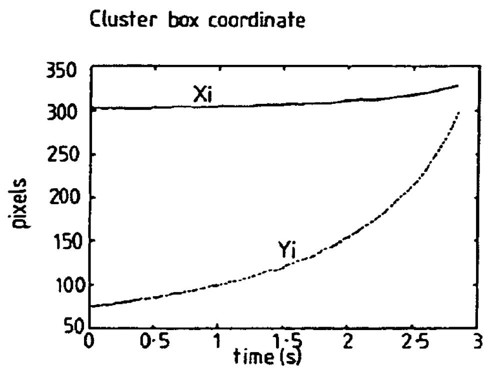

FIG. 26

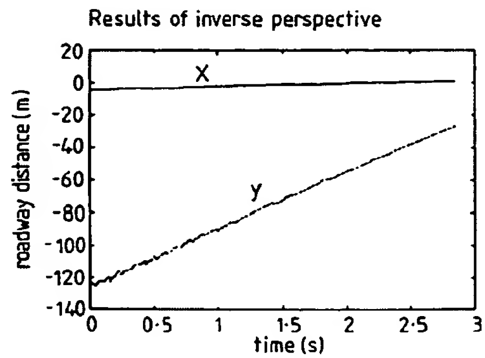

FIG. 27

Figure 26 shows the change over time of the coordinates of a vehicle in the detector camera image plane. This shows the natural effect of the vehicle moving from top to bottom in the image more quickly as it gets closer to the detector camera. Figure 27 shows the calculated distance of the vehicle based on the pixel locations of Figure 26. The location numbers in Figure 27 are generated by applying the pixel scaling factors and other conversion elements of the conversion equation to the pixel coordinates of Figure 26.

The cited passage concludes as follows:

> The time at which the vehicle 302 will reach the acquisition line 22, and the future location of the vehicle 302, need to be predicted, due to latency in the system. Considerable latency exists between a trigger request and image acquisition via the acquisition camera 8, and additional latency is caused by pixel transfer, image processing pipeline delay and software processing delay. The information obtained on the basis of the images required by the detection camera 6 provide a delayed representation of the actual vehicle position, and therefore it is necessary to estimate the future position and speed of the vehicle 302.

This portion confirms that the distance measurements generated by the conversion equation are used to track the vehicle and trigger the operation of the main camera.

Applicant sincerely believes that the cited passage does not teach the conversion of the resolution of images so that their resolutions coincide. Rather,

it teaches converting the coordinates of a vehicle in the detector image into real-world coordinates. This is done using a single detector camera. The main camera is not involved in determining how far away the vehicle is, and no conversion is done to make the resolution of the detector camera coincide with the resolution of the main camera. There is no reason to perform any conversion on the detection camera image to alter its pixel units in any way with reference to those of the main camera image, and no such conversion is done.

It is therefore believed that the present claims are allowable over the cited references. All of the pending independent claims recite the feature of converting the resolution of images formed by cameras having different resolutions or different fields of view so that the resolutions of the images coincide after conversion, and then using those images to perform distance measurements through triangulation. Applicant's admitted prior art teaches triangulation using two cameras but does not teach to using cameras that have different resolutions or different fields of view, and then converting the resolutions of the images of those cameras so that they coincide before performing triangulation. Auty teaches a system that involves two cameras and performs a distance measurement, but the system performs the distance measurement using only one camera, and does not (and has no need to) convert the image from that camera so that both cameras' resolutions coincide. Therefore the combined teachings lack the features required by the claims. Further, Auty's measurement is done using a single camera in manner explained above, rather than by triangulation, and one would not be motivated to use features of Auty in a triangulation system since they would simply not have any application there. Auty does not provide any solution to the problem of how to perform triangulation when the cameras that are being used have different fields of view or different resolutions. Lacking such teaching, Auty cannot suggest the features recited in the present claims.

The foregoing remarks address all bases for rejection and show that all claims are allowable. The undersigned will contact the examiner to set up a telephone conference to discuss these points further.

Respectfully submitted,

Date: 10 December 2003 By Ronald Coslick

FOLEY & LARDNER
Washington Harbour
3000 K Street, N.W., Suite 500
Washington, D.C. 20007-5109
Telephone: (202) 672-5407
Facsimile: (202) 672-5399

Ronald Coslick
Attorney for Applicant
Registration No. 36,489